Regulatory Toxicology and Pharmacology **32,** 56–67 (2000)
doi:10.1006/rtph.2000.1399, available online at http://www.idealibrary.com on IDEAL®

# Concordance of the Toxicity of Pharmaceuticals in Humans and in Animals

Harry Olson,[1] Graham Betton,[2] Denise Robinson,[3] Karluss Thomas,[3] Alastair Monro,[1] Gerald Kolaja,[4] Patrick Lilly,[5] James Sanders,[6] Glenn Sipes,[7] William Bracken,[8] Michael Dorato,[9] Koen Van Deun,[10] Peter Smith,[11] Bruce Berger,[12] and Allen Heller[13]

[1]*Pfizer Inc., Groton, Connecticut;* [2]*AstraZeneca Pharmaceuticals, Macclesfield, England;* [3]*ILSI-HESI, Washington, DC, 20036;* [4]*Pharmacia & UpJohn, Kalamazoo, Michigan;* [5]*Boehringer Ingelheim Pharmaceuticals, Ridgefield, Connecticut;* [6]*Rhone-Poulenc Rorer, Collegeville, Pennsylvania;* [7]*University of Arizona, Tucson, Arizona;* [8]*Abbott Laboratories, Abbott Park, Illinois;* [9]*Eli Lilly and Co., Greenfield, Indiana;* [10]*Janssen Research Foundation, Beerse, Belgium;* [11]*Monsanto-Searle Laboratories, Skokie, Illinois;* [12]*Sanofi-Synthelabo, Inc., Malvern, Pennsylvania; and* [13]*Bayer Corporation, West Haven, Connecticut*

Received January 22, 2000

**This report summarizes the results of a multinational pharmaceutical company survey and the outcome of an International Life Sciences Institute (ILSI) Workshop (April 1999), which served to better understand concordance of the toxicity of pharmaceuticals observed in humans with that observed in experimental animals. The Workshop included representatives from academia, the multinational pharmaceutical industry, and international regulatory scientists. The main aim of this project was to examine the strengths and weaknesses of animal studies to predict human toxicity (HT). The database was developed from a survey which covered only those compounds where HTs were identified during clinical development of new pharmaceuticals, determining whether animal toxicity studies identified concordant target organ toxicities in humans. Data collected included codified compounds, therapeutic category, the HT organ system affected, and the species and duration of studies in which the corresponding HT was either first identified or not observed. This survey includes input from 12 pharmaceutical companies with data compiled from 150 compounds with 221 HT events reported. Multiple HTs were reported in 47 cases. The results showed the true positive HT concordance rate of 71% for rodent and nonrodent species, with nonrodents alone being predictive for 63% of HTs and rodents alone for 43%. The highest incidence of overall concordance was seen in hematological, gastrointestinal, and cardiovascular HTs, and the least was seen in cutaneous HT. Where animal models, in one or more species, identified concordant HT, 94% were first observed in studies of 1 month or less in duration. These survey results support the value of *in vivo* toxicology studies to predict for many significant HTs associated with pharmaceuticals and have helped to identify HT categories that may benefit from improved methods.** 

## INTRODUCTION

A vitally important theme in toxicology is the search for and the assessment of *in vitro* and *in vivo* models that are predictive for adverse effects in humans exposed to chemicals. The conduct of toxicology studies in laboratory animals is driven by experience, historical precedence, and governmental requirements, and the results of these studies usually, and reasonably, lead to restrictions on the use, or method of use, of the chemicals concerned. Such a process must be based on the assumption that the current choice of animal models and the design of the studies are truly predictive of human hazard. The reliability of this assumption has far-reaching repercussions in terms of the potential for inappropriate use of animals and the unnecessary deprivation of, or restrictions in the use of, valuable chemicals including pharmaceuticals. Identification of any weaknesses in the assumption could lead to revisions of existing regulations and stimulate the search for better methods for the safety evaluation of chemicals in the future.

There have been relatively few attempts to methodically assess the correlation between the toxicity caused by chemicals in animals and in humans. This is not surprising, given that the toxicity of many chemicals observed in humans is after accidental exposure, the quantitative details of which in terms of duration and intensity are often not known. Chemicals, which are components of the diet, either macro- or micro-, are more susceptible to evaluation of their toxicity in animals and in humans, provided that the means to carry out epidemiological studies are available. However, a rich source of relevant information is pharmaceutical chemicals. For these, the human exposure is controlled and measured accurately. In addition, clinical studies of drugs employ systematic clinical examinations and

0273-2300/00 $35.00

tests of organ function, aimed specifically at the detection of adverse effects.

There are few published analyses of comparative animal–human toxicity data on pharmaceuticals, with progress presumably inhibited by the perceived confidential nature of such data. The first brave foray into this area was by Litchfield when he analyzed an array of toxicities in rats, dogs, and humans for six diverse drugs being developed by his company (Litchfield, 1962). He reported that toxicities that occurred in rats only were rarely observed in humans and those in dogs only occurred slightly more frequently in humans, while those that occurred in both rats and dogs showed about a 70% concordance with humans.

Cytotoxic anticancer agents, by their very nature, tend to cause much toxicity in humans and several groups have examined the extent to which toxicities seen in humans can be predicted from animal data for these drugs (Owens, 1962; Schein *et al.,* 1970; Rozencweig *et al.,* 1981). Generally, these drugs caused qualitatively similar toxicity in animals and in humans, with data from dogs predicting gastrointestinal toxicity in humans particularly well and data from dogs and monkeys grossly overpredicting hepatic and renal toxicity. Rozencweig *et al.* (1981) warned that the predictability of such data is highly dependent on the prevalence of the particular human toxicity, with rare toxicities being essentially unpredictable from animal data.

Clinical toxicity data for more diverse types of drugs have also been the subject of several workshops and overviews (Lawrence *et al.,* 1984; Fletcher, 1987; Lumley and Walker, 1990; Parkinson *et al.,* 1994). In one small series in which the toxicity in clinical trials led to the termination of drug development, it was found that in 16/24 (67%) cases the toxicity was not predicted in animals (Lumley, 1990). In another analysis, 39/91 (43%) clinical toxicities (from 64 marketed drugs) were not predicted from animal studies (Igarashi, 1994). This latter publication forms part of the largest data set known to us, that of the Japanese Pharmaceutical Manufacturers Association (JPMA, 1994). This was derived from the literature (as distinct from questionnaire-derived data) and refers to data from 139 drugs approved in Japan from 1987 to 1991. The animal toxicity data are drawn from 468 repeated-dose studies, mainly in rats and dogs but with a few studies in mice and monkeys. No indication was given about the importance of the clinical toxicity, e.g., whether it was trivial or whether it led to a restriction in the use of the drugs. There were few correlations across species with, overall, the best predictivity being for cardiovascular events, and the poorest for cutaneous and hypersensitivity phenomena. Despite its relatively high incidence in all species, hepatobiliary toxicity in humans was surprisingly poorly predicted from animal studies. The JPMA also conducted an enterprising review of the extent to which general pharmacology studies were useful in predicting adverse effects in humans (Igarashi *et al.,* 1995). A total of 141 drugs were reviewed and the analysis showed considerable value in tests of spontaneous locomotor activity in mice, gastrointestinal transit time in mice, gastric secretion in rats, and urinary retention and sodium excretion in rats.

Two reviews addressed those drug cases where the clinical toxicity was so severe as to lead to withdrawal from marketing in the approximate period 1960–1990 (Heywoood, 1990; Spriet-Pourra and Auriche, 1994). In one report only 4 of 24 cases were predictable from animal data; in the other report, only 6 of 114 clinical toxicities had animal correlates. Such a poor correspondence is not surprising, given that these late-onset phenomena once on the market are usually idiosyncratic in nature, i.e., of very low incidence, not dose-related, and apparently not related to the pharmacology of the compound.

A knowledge of pharmacology in various species, including humans, tells us that species can differ markedly in their response to pharmacological agents. Indeed, it has been reported that 29% of withdrawals of drugs from development are attributable to an inappropriate (e.g., lack of efficacy or selectivity) pharmacological response (Prentis *et al.,* 1988). Against that background, one would expect diverse species responses to many toxic stimuli. Several reviews (Oser, 1981; Calabrese, 1984, 1987; Garratini, 1985; Zbinden, 1993) summarize and discuss the many differences in anatomy, physiology, or biochemistry between laboratory animals and humans and these can provide useful points of reference for anticipating whether a particular chemical is likely to show a similar response in an animal species and in humans. In addition to the fundamental differences between species in biological response, Zbinden (1991) cautioned against too great an expectation from animal toxicology studies for a host of reasons inherent in the designs of such studies and of clinical trials (see Table 1).

In a symposium that addressed the question of the relevance of animal toxicology studies for humans, several contributors urged the pharmaceutical industry to collaborate and pool its data on toxicity of drugs in development with a view to drawing broadly based conclusions (Brimblecombe, 1990; McLean, 1990). In this paper we report the first product of just such a collaboration. Twelve companies provided codified data to the International Life Sciences Institute (ILSI) who coordinated the compilation and analysis of the data. The primary objective was to examine how well toxicities seen in preclinical animal studies would predict actual human toxicities for a number of specific target organs using a database of existing information. To a lesser degree, the symposium also sought to better understand the duration of dosing required in animals

**TABLE 1**
**Some Differences between Animals and Humans Critical to Prediction of Toxicity**

| | Animals | Man |
|---|---|---|
| Subjects | | |
| Number | Large groups | Individuals |
| Age | Young adult | All ages |
| State of health | Healthy | Usually sick |
| Genetic background | Homogeneous | Heterogeneous |
| Doses | | |
| Magnitude | Therapeutic to toxic | Therapeutic |
| Schedule | Usually once daily | Therapeutic optimum |
| Circumstances | | |
| Housing | Uniform, optimal | Variable |
| Nutrition | Uniform, optimal | Variable |
| Concomitant therapy | Never | Frequent |
| Diagnostic procedures | | |
| Verbal contact | None | Intensive |
| Physical exam | Limited | Extensive |
| Clinical lab | Limited, standardized | Individualized |
| Timing | Predetermined | Individualized |
| Autopsy | Always | Exceptional |
| Histopathology | Extensive | Exceptional |

to reveal the same toxicity in man where the same toxicity was seen in animals and man.

## APPROACH TO DATA COLLECTION AND METHODS OF ANALYSIS

Although a considerable effort was made to collect data that would enable a direct comparison of animal and human toxicity, it was recognized from the outset that the data could not answer completely the question of how well animal studies predict overall the responses of humans. To achieve this would require information on all four boxes in Fig. 1, and this was not practicable at this stage. The magnitude of the data collection effort that this would require was considered impractical at this stage. The present analysis is a first step, in which data have been collected pertaining only to the left column of Fig. 1: true positives and false negatives. By definition, therefore the database only contains compounds studied in humans (and not on those that never reached humans because they were considered too toxic in animals or were withdrawn for reasons unrelated to toxicity). Despite this limitation, it was deemed useful to proceed in the expectation that any conclusions that emerged would address some of the key questions and focus attention on some of the strengths and weaknesses of animal studies.

The approach chosen was to collect and collate toxicities in humans (HTs) associated with exposure to pharmaceuticals. Since most data resided in the archives of the larger companies with long histories of drug development, these companies were approached to participate. The companies were asked to tabulate HTs that had occurred during drug development. No restriction was placed on the time frame from which the data were drawn, except that it was requested that participants provide data from compounds in clinical development over a consecutive period of years, to avoid any bias from selected data sets.

A working party of clinicians from participating companies developed criteria for "significant" HTs to be included in the analysis. For inclusion a HT (a) had to be responsible for termination of development, (b) had to have resulted in a limitation of the dosage, (c) had to have required drug level monitoring and perhaps dose adjustment, or (d) had to have restricted the target patient population. The HT threshold of severity could be modulated by the compound's therapeutic class (e.g., anticancer vs anti-inflammatory drugs). In this way, the myriad of lesser "side effects" that always accompany new drug development but are not sufficient to restrict development were excluded. The judgments of the contributing industrial clinicians were final as to the validity of including a compound. The clinical trial phase when the HT was first detected and whether HT was considered to be pharmacology-related was recorded. HTs were categorized by organ system and detailed symptoms according to standard nomenclature (COSTART, National Technical Information Service, 1999).

In a subsequent step the company toxicologist examined the reports of toxicology studies in animals for the compounds that met the criteria of HT. Data examined included clinical signs, physiological measurements, hematology and clinical chemistry assays, and histopathology evaluations from rodent and nonrodent toxicology or safety pharmacology studies. A toxicity correlation was considered to be positive if the same target organ was involved in humans and in animals in the judgment of the company clinicians and the toxicologists. In the case of a positive correlation, the data provided were the rodent and nonrodent species tested and the duration of exposure (often the same as the duration of the study) at which the toxicity in question was first observed. In the event of lack of correlation, the data provided included all species tested, the longest duration of dosing in each species tested, whether dose-limiting toxicity was achieved, and whether there

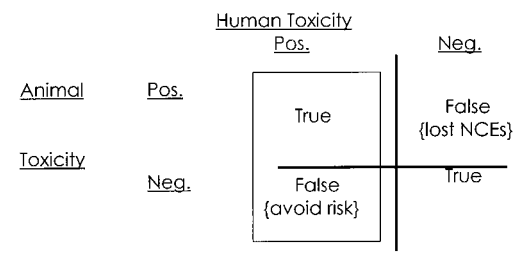


**FIG. 1.** Predictivity of animal toxicity data.

was a qualitative similarity in the patterns of metabolism in humans and in animals that were available.

All data were provided to the ILSI staff who entered them in codified fashion to maintain confidentiality. A central database of 221 HT examples from 150 compounds was developed for analysis. Three academic experts (Dr. G. Sipes of University of Arizona, Dr. R. Bain of George Washington University, and Dr. R. Abernethy of Georgetown University) advised on the study conduct of the analyses and interpretation of the data.

An ILSI Workshop was held in Virginia in April 1999, involving academic scientists, government regulators, and industry scientists, to review the data. At the Workshop, the participants were divided into groups to discuss the six principal types of HT reported: hepatic, neurological, cardiovascular, hematological, gastrointestinal, and hypersensitivity. These groups were asked to address the same series of generic questions on the data for their respective toxicities. Finally, a panel of experts was convened to assess the value and utility of the database and to make recommendations about the continuation of the project in the future.

## RESULTS (PART 1)

Results are presented in two parts: first, overall analyses of the total database; and second, the answers of the breakout groups to the generic questions.

### Overall Analysis of Findings (Total Database)

Preliminary results of analysis of the incomplete database have been reported previously (Olson, 1998).

#### *Distribution of Human Toxicities by Therapeutic Class*

Overall, a total of some 221 separate cases of compounds associated with significant human toxicity were recorded. A total of 150 compounds contributed to this series with multiple HTs being recorded in 47 cases. The distribution of the therapeutic class of compounds studied is shown in Table 2.

For each of the compounds in the database, information was collected regarding the route of administration. The routes of administration employed in humans were 168 HTs by oral, 52 intravenous, 7 by inhalation, and 2 dermal, with two routes of administration being used in the case of 9 HTs.

The therapeutic classes showed significant variations in their COSTART organ-system-associated HT profile as shown in Table 3. Detailed analyses of signs and symptoms within COSTART groupings are addressed in the following sections.

The rate of project termination for various HTs was highest for (in order) urogenital, cutaneous, hepatic, and cardiovascular HTs and, by therapeutic class, highest for anti-inflammatory, antiviral, endocrine, and respiratory therapeutic classes (Table 3).

**TABLE 2**
**Distribution of Compounds by Therapeutic Class**

| Therapeutic class | No. of compounds |
|---|---|
| Anticancer | 14 |
| Anti-infection | 21 |
| Anti-inflammatory | 15 |
| Antiviral | 8 |
| Cardiovascular | 16 |
| Endocrine | 10 |
| Gastrointestinal | 9 |
| Hematology | 1 |
| Immunology | 2 |
| Impotence | 2 |
| Metabolism | 5 |
| Neurologic | 31 |
| Renal | 2 |
| Respiratory | 13 |
| Trauma | 1 |
| Total | 150 |

#### *Relationship to Dosing Duration and Clinical Trial Phase*

The time of first onset of HTs according to the clinical trial phase was analyzed according to HT class. Overall, over half of HTs were first manifest in Phase I trials. HTs seen after single-dose administration to man numbered 62 cases with 158 cases seen following multiple doses (remainder unspecified). Classes of HTs detected with frequency in Phases II and III were cutaneous and hepatic types (Table 4).

The survey also recorded the frequency of development project termination. In those instances where the HT led to project termination, 39% were terminated in Phase I, 43% were terminated in Phase II, and 10% were terminated in Phase III.

Only four HTs were considered to be idiosyncratic in nature, two cases of rash (one in Phase I, one in Phase II) and two cases of thrombocytopenia in Phase II.

#### *Pharmacologic Basis of Human Toxicities*

The characterization of HTs as being related to the primary pharmacological activity of the drug is given according to HT class in Fig. 2.

The overall distribution of pharmacological HTs according to clinical trial phase was 35% in Phase I, 39% in Phase II, and 43% in Phase III.

#### *Concordance by Animal Models*

*Concordance by one or more species: Overall and by HT.* Overall, the true positive concordance rate (sensitivity) was 70% for one or more preclinical animal

**TABLE 3**
**Frequency of Human Toxicities According to Therapeutic Class and Percentage of Terminations**

| Therapeutic class | BCH | CUT | HEP/LFT | CV/ECG | END | NRL | HEM | GI | MSK | REPRO | URN | OTH | TOTAL | %TERMN |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Anticancer | 1 | 1 | 2 | 1 | 1 | 5 | 6 | 3 | 0 | 0 | 2 | 3 | *25* | 20 |
| Anti-infection | 0 | 3 | 6 | 2 | 0 | 6 | 2 | 10 | 3 | 0 | 1 | 5 | *38* | 37 |
| Anti-inflammatory | 0 | 4 | 2 | 2 | 0 | 6 | 0 | 6 | 0 | 0 | 2 | 0 | *22* | 55 |
| Antiviral | 0 | 1 | 3 | 0 | 0 | 1 | 1 | 2 | 0 | 0 | 3 | 0 | *11* | 54 |
| Cardiovascular | 0 | 0 | 1 | 11 | 0 | 4 | 1 | 1 | 0 | 0 | 0 | 0 | *18* | 39 |
| Endocrine | 0 | 2 | 3 | 2 | 2 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | *12* | 50 |
| Gastrointestinal | 0 | 0 | 2 | 2 | 0 | 4 | 0 | 3 | 0 | 0 | 0 | 3 | *14* | 36 |
| Hematology | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | *1* | 0 |
| Immunology | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | *2* | 0 |
| Impotence | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | *4* | 0 |
| Metabolism | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | *5* | 20 |
| Neurologic | 0 | 2 | 5 | 8 | 0 | 18 | 0 | 11 | 1 | 0 | 1 | 2 | *48* | 33 |
| Renal | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | *2* | 100 |
| Respiratory | 0 | 1 | 6 | 2 | 2 | 1 | 0 | 0 | 0 | 1 | 1 | 3 | *17* | 47 |
| Trauma | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | *2* | 0 |
| Total | 1 | 14 | 31 | 36 | 5 | 49 | 11 | 41 | 4 | 1 | 10 | 18 | 221 | 37 |
| % Terminated by HT | 0 | 64 | 55 | 47 | 40 | 35 | 27 | 10 | 25 | 100 | 70 | 6 | | |

*Note.* BCH, biochemical; CUT, cutaneous; END, endocrine; GI, gastrointestinal; HEM, hematologic; HEP/LFT, hepatobiliary and liver function test abnormalities; MSK, musculoskeletal; NRL, neurological; REPRO, reproductive; URN, urinary; OTH, other.

model species (either in safety pharmacology or in safety toxicology) showing target organ toxicity in the same organ system as the HT (Fig. 3). For the remaining 30% of HT there was no relationship between toxicities seen in animals and those observed in humans. Concordance was seen in 63% of nonrodent studies (primarily the dog) and 43% of rodent studies (primarily the rat). There was considerable overlap in toxicology with 36% of HTs being concordance with two species (i.e., a rodent and a nonrodent) with concordance by only one species occurring in the nonrodent (27% of HTs) and the rodent (7%).

The ratio of positive concordance versus nonconcordance by rodent and nonrodent species is shown in Fig. 4. The total incidence of usage of each species (concordance and nonconcordance) was incorporated and showed the nonrodent species of dog and primate to have a higher frequency of positive concordance than did rodents.

Concordance varied significantly according to the human target organ system affected as shown in Fig. 5.

The best concordance was for hematological, gastrointestinal, and cardiovascular toxicities and the least was for cutaneous toxicity. The proportional contribution of nonrodent versus rodent models to concordance for given types of HTs is discussed below for the main organ system HTs observed. There were marked differences in the relative contribution of nonrodent and rodent toxicology species to concordance depending on the HT category (Table 5).

Analysis of overall prediction rates according to clinical trial phase when the HT was first observed showed a slightly higher frequency in Phase I onset HTs (75%) compared to Phase II (58%) and Phase III (52%).

*Concordance by therapeutic class.* Since certain therapeutic classes were prone to expression of given types of HTs, e.g., hematotoxicity with anticancer agents, the variation in concordance rates according to HT class could indirectly influence the concordance rates for certain therapeutic classes. This is reflected in Fig. 6.

*Time to first appearance of concordant animal toxicity.* Where the animal model(s) were successful in predicting for a given HT, the survey requested the earliest time at which the relevant animal toxicity was

**TABLE 4**
**Distribution of Clinical Trial Phase Time of First Onset by HT Class**

| Therapeutic class: Phase | BCH | CUT | CV/ECG | END | GI | HEP/LFT | HEM | MSK | REPRO | NRL | URN | OTH | TOTAL | TERMN | % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | 1 | 8 | 23 | 1 | 30 | 14 | 7 | 0 | 1 | 29 | 6 | 15 | 135 | 52 | 39 |
| II | 0 | 6 | 10 | 4 | 7 | 13 | 4 | 2 | 0 | 14 | 2 | 3 | 65 | 28 | 43 |
| III | 0 | 0 | 3 | 0 | 4 | 4 | 0 | 2 | 0 | 6 | 2 | 0 | 21 | 2 | 10 |

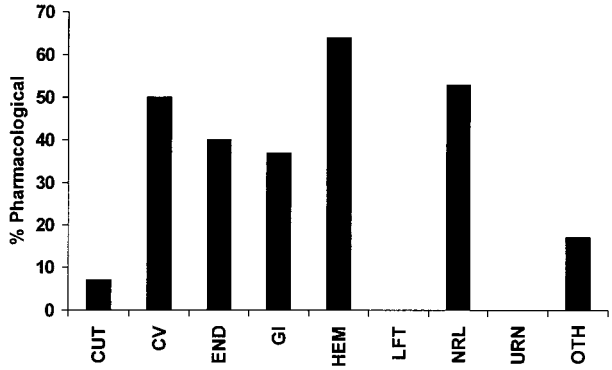


**FIG. 2.** Percentage of HTs judged to be related to the primary pharmacology.

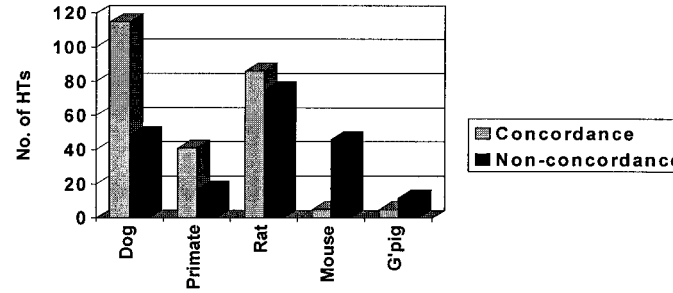


**FIG. 4.** Concordance rates versus species.

observed. Overall, 94% of animal target organ toxicities correlated with HTs were first observed in studies less than or equal to 1 month in duration (Fig. 7). A large proportion of animal toxicities was observed following single dose administration; 25% of these observations were from safety pharmacology rather than toxicity studies.

*False Negative Prediction by Animal Models*

*Duration of animal studies where HT was not concordant (false negatives).* Given the high rate of detection by animals of HTs in studies of 1 month duration or less, it is possible that failure to detect HTs may have resulted from insufficient duration of exposure in animal models. Analysis of the longest duration of studies conducted in animals shows that overall 73% of HT cases had been studied in animals for 2 months or more (Fig. 8).

Since 61% of HTs first occurred during Phase I clinical trials when chronic toxicology programs may not have been completed, the duration of animal testing on this subset was also analyzed: 45% of these Phase I onset cases had one or more toxicology species tested for 2 months or longer.

*Achievement of limiting toxicity in animals in false negative cases.* Since interspecies differences in exposure (toxicokinetic data were not collected in the survey) may have resulted in failure to demonstrate relevant target organ toxicity, the proportion of animal studies achieving dose limiting toxicity [up to maximum tolerated dose (MTD) in some cases] was analyzed. For cases where HTs were not predicted, 91% of rodent and 90% of nonrodent toxicology studies were judged to have been performed at limiting doses. Hence, insufficient exposure of animals to drug alone could not account generally for the 30% false negative rate.

*Correlation of animal metabolite profile with man in false negative cases.* Using a qualitative judgment of whether the main human metabolites were present in one or more animal toxicology species (data available in 29 of 63 false negative cases), animal metabolism profiles were considered to correlate with that of man in 86% of cases. Therefore, metabolic differences between animals and man alone probably do not explain the false negative cases. Taking both concordant and nonconcordant cases overall where comparative metabolism data were available, there was a 89% animal: human metabolite correlation rate.

*Correlation of animal pharmacological responsiveness in false negative cases.* Since approximately 40% of HTs were evaluated as being pharmacology-related, pharmacological unresponsiveness in the animal species could result in false negative prediction. Taking false negative prediction cases, the animal species used as models were, in one or more species, considered to be

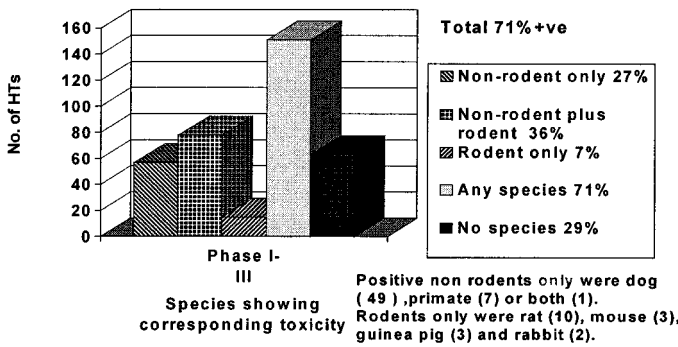


**FIG. 3.** Concordance of human toxicity from animals.

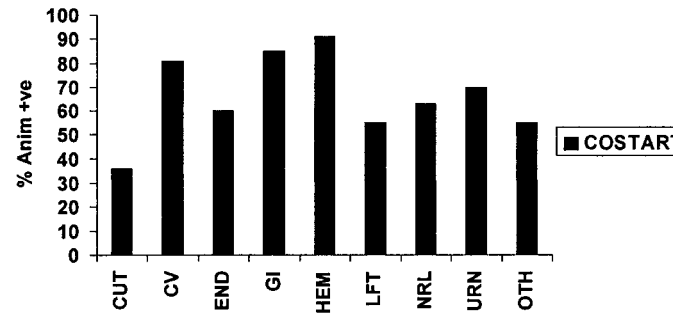


**FIG. 5.** Animal concordance by HT COSTART category.

**TABLE 5**
**Animal Species Concordant Number of Cases for Frequent HT COSTART Systems**

| COSTART: Animal correlates | CUT | CV/ECG | GI | HEM | HEP/LFT | NRL |
|---|---|---|---|---|---|---|
| Yes | 5 | 29 | 35 | 10 | 17 | 31 |
| No | 9 | 5 | 6 | 1 | 14 | 13 |
| Nonrodent only | 1 | 16 | 16 | 0 | 7 | 11 |
| Rod and nonrodent | 1 | 12 | 18 | 10 | 8 | 17 |
| Rodent only | 3 | 1 | 1 | 0 | 2 | 3 |

pharmacologically responsive in 63% of cases. Taking only those false negative cases of pharmacology-related HT prediction, 85% of animal models were pharmacologically responsive. Pharmacological unresponsiveness of the animal models therefore cannot alone account for the false negative rate. Taking both the true positive and the false negative cases together, 69% of animal species used were pharmacologically responsive with 87% being responsive for pharmacology-related HTs overall.

## RESULTS (PART 2)

This section is a compilation of the responses of the six breakout groups at the ILSI Health and Environmental Science Institute's (HESI) Workshop (April 1999) which focused on the six principal types of HT. Each breakout group addressed the four main questions, listed below. These results include group comments which are similar; specific remarks from individual breakout groups are so noted.

*Q.1. Evaluate the Database Generally in the Context of the Breakout Group's Specific Endpoints and Comment on: (i) Any Animal–Human Toxicity Correlations That Can Be Made; (ii) Whether the HT Was Related to the Therapeutic Class or the Known Primary Pharmacology of the Compound*

Most groups were critical of the database to the extent that it often lacked specific detail regarding the exact nature of the HT, including its incidence in patients, severity, and time to onset. Uncertainties in nomenclature, application of the COSTART terminology, and whether the HT was multifaceted were also perceived to be problems. The *liver* group would like to have known whether the HT was hepatocellular or biliary cell injury, jaundice, or fulminant liver failure. The *hematology* group would like to have had more information on the type of cells affected and to have known if neutropenia existed alongside other hematotoxicity. The *cardiovascular* group would have been interested to know whether hypotension was accompanied by tachycardia.

The tentative conclusions regarding the incidence of HT and of termination by clinical phase (Results (Part 1), "Relationship to Dosing Duration and Clinical Trial Phase") must be tempered with caution. The HT data supplied to the database did not distinguish between the severity of the HT and its incidence. Thus a HT first observed in Phase III may have been the result either of a longer drug exposure or of a greater number of patients being in the trial than in earlier clinical phases. For example, termination may have been caused by 1% incidence of a severe HT or a 50% incidence of a mild HT.

Several groups noted limitations (already alluded to, see Table 1) inherent in such a human–animal comparison exercise. Thus, failure to distinguish between toxic signs and symptoms, and especially nausea and many of the subjective neurologic HTs (e.g., headache, hallucinations), emphasized the need to limit expectations for prediction of such HTs. Interpretation was

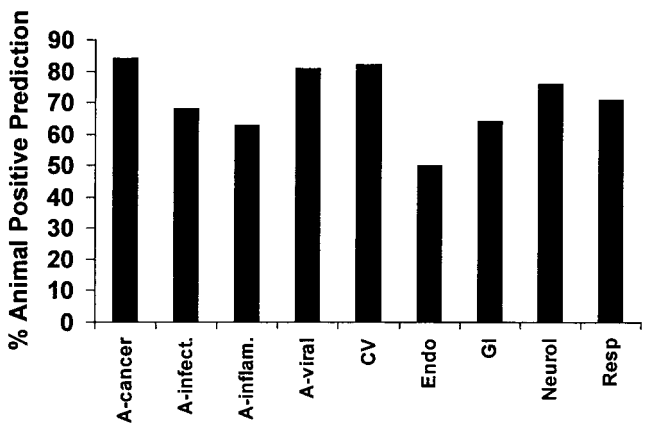


**FIG. 6.** Preclinical concordance for HTs by therapeutic class.

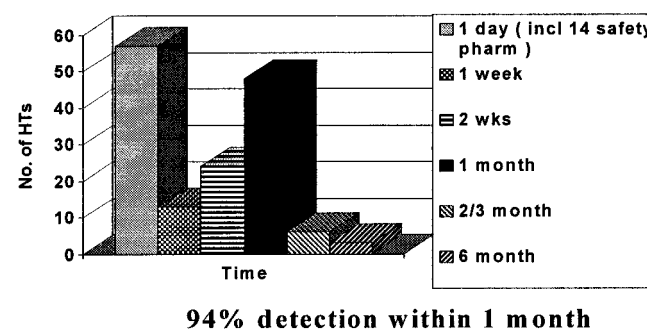


**FIG. 7.** Time to first detection of relevant toxicity in animals.

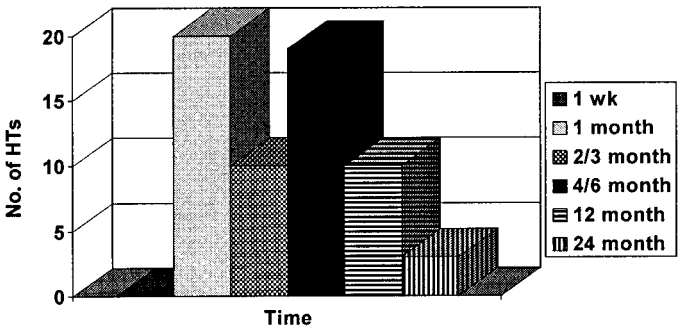


**FIG. 8.** Longest duration of nonconcordant toxicity studies ($N$ = 73).

also impeded by uncertainty over what specific evidence of toxicity in animals was considered to have predicted each HT. In general, it had to be assumed that any disturbance of the relevant organ system in animals was judged to be a positive correlation. For example, diarrhea in animals might correspond in humans to abdominal pain or nausea; excitation in animals to, say, dizziness in humans; and so on.

The *hypersensitivity* group observed that tests for such activity were rarely included in preclinical toxicology packages and thus correlations would not be expected with the data routinely obtained from animals. Also, while most HTs in this category were cutaneous but there was a notable data gap with few compounds applied via the dermal route. Also a proven immunological basis for the symptoms of rash was not evident in all cases.

The discrepancies between the incidence of each HT and the frequency with which that HT led to termination in clinical trials were interesting. It was recognized by several groups that the decision to terminate or continue in clinical trial will, of necessity, depend on the therapeutic ratio and the seriousness of the condition being treated and also the stage of clinical development and investment with few Phase III HTs resulting in project termination.

*Liver* toxicity was only the fourth most frequent HT (Table 3), yet it led to the second highest termination rate. There was also less concordance between animal and human toxicity with regard to liver function, despite liver toxicity being common in such studies. There was no relation between liver HTs and therapeutic class. This poor correlation prompted the question of whether a partial explanation might lie in the assumption that the same biomarkers were as appropriate in humans as in animals. The first signs of unwanted effects on the hepatobiliary system in humans are usually rises in circulating aminotransferase enzyme levels. The JPMA study (Igarashi, 1994) reported that these enzymes are relatively insensitive markers of liver toxicity in animals, but that the correlation with human toxicity was much better if one incorporated histopathology data from animals (also included in the present survey).

*Neurologic* HTs were the most common category (22%) and they occurred disproportionately more frequently for neurologic drugs. It was noted that whereas the correlation with animal studies was independent of the therapeutic class of drug, these HTs led to termination of only 17% of neurologic drugs but to 45% of nonneurologic drugs. Such HTs are apparently more acceptable in neurologic drugs, this obviously depending to some extent on the concurrent therapeutic benefit. Correlations with findings in animals were much better for nonrodents than for rodents. The cases in rodents were all peripheral neuropathies in anticancer agents.

*Cardiovascular* HTs had a high rate of concordance (80%) with animal studies. Most were caused by cardiovascular drugs (Table 2) and interpretable in terms of their primary pharmacological activity. The main categories were: rhythm changes (tachycardia, bradycardia), ECG abnormalities, hypotension, and vasodilatation. While data from rodents were indicative of a potential for hypotension, rodents contributed no information that was not available from nonrodents although the rabbit detected an ECG HT case. By contrast, 28/36 of the HTs were observed in nonrodents (25 in dogs, 6 in monkeys).

*Hematologic* HTs correlated in high degree (91%) with animal findings, with rodents and nonrodents both being responsive. Termination due to this HT was relatively low (27%). This HT was strongly associated with anticancer and anti-infective agents, with only two compounds coming from other therapeutic classes.

*Gastrointestinal* HTs overall correlated well (85%) with animal findings, especially in nonrodents. This HT was associated primarily with anticancer, anti-infective, and anti-inflammatory agents, all known to provoke such toxicity by well-understood pharmacological mechanisms. Gastrointestinal HTs were the second most common category, yet had the lowest rate of termination. One might say that these HTs are apparently often just regarded as "nuisance" effects.

*Hypersensitivity* HTs (essentially various cutaneous reactions) were too few to allow conclusions about relationship to therapeutic class or pharmacological mode of action, in particular whether the HTs were immunological or not in character. Nevertheless, the high rate of termination (64%), the highest in the database, highlighted the need to improve preclinical testing methodology in this area (*vide infra*).

### *Q.2. What Duration of Dosing in Animals Was Sufficient to Reveal the Toxicity That Corresponded to a HT?*

Several groups pointed out that the minimum times given in Fig. 7 may well be overestimates for many tox-

icities, i.e., those dependent on time of sampling for biochemical or hematological tests, on time of functional investigation, or on time of termination for histopathological examination. Even so, the data in Fig. 7 are striking in that 57 (38%) of the relevant toxicities were observed on Day 1, these being primarily cardiovascular (16), gastrointestinal (15), and neurological (12) phenomena. Liver toxicity was never reported after a single dose. Most of the single dose observations for cardiovascular events came from specific safety pharmacology studies; the others came mainly from clinical observations on the first day of a multidose study.

The nine HTs that required over 1 month to manifest themselves in animals showed no pattern. They comprised one or two agents from each of the HT categories.

*Q.3. Which HTs Were Not Detected by the Animal Studies?*

The *neurologic* group inevitably found many examples of nondetection (or perhaps "not known") because of the need to communicate symptoms such as headache, dizziness, etc. The only objective sign not predicted (seizure) was associated with a difference in *metabolism* between humans and the animals tested.

The *cardiovascular* group found a case of myocardial infarction not predicted from animal studies, perhaps because of a lack of an animal counterpart. Two cases of hypotension were not observed in animals: in one the MTD was not achieved in animal studies; in the other a clear difference in animal/human metabolism was noted.

The *hematologic* HT that sometimes escaped detection was thrombocytopenia: two cases, even after 6-month studies in animals. The panel suggested that this toxicity could have been detected by nonstandard methodology.

The only *gastrointestinal* HT that did not correlate with animal studies was nausea. This is not surprising, given its subjective nature and the uncertainty regarding whether it should be classified here or as a neurologic HT.

No cutaneous (classified here as *hypersensitivity*) HTs were observed using the standard animal studies conducted, although phototoxicity did correlate well with the response of guinea pigs in special tests.

*Q.4. Identify Novel and/or Available Technologies That Could Be Included in Existing Animal Studies to Address Deficiencies in Identifying Target Toxicity, and Identify Testing Strategies—Including Additional Studies—That Could Be Implemented to Improve for Screening Compounds into Development*

The major recommendations (not in any particular order) were the following:

a. It was suspected that some of the poor correlations may have been due to apparent design deficiencies in animal studies. Thus, the hematology breakout group recommended the use of toxicokinetic and tissue distribution studies to lead to better timing of blood sampling and function tests. Safety pharmacology studies should be subject to the same considerations, as recommended previously by Jorkasky (1998).

It was recognized that interspecies differences in pharmacokinetics are unlikely to underlie many of the noncorrelations in this database, given the manifestly different pharmacological responses of species and the fact that almost all the animal studies were carried out at toxicity-limited dose levels. Nevertheless, the quantitative aspect cannot be completely disregarded. There must be certain pharmacological mechanisms that allow approximate allometric relationships in dose–responses across species. If established, this would give a perspective on toxicity observed in animals only at enormous multiples of the expected clinical dose and would provide guidance to a clinician planning to increase the dose level of a drug in a clinical trial.

b. The *cardiovascular* and *neurologic* groups, cognizant of the importance of safety pharmacology studies for these HTs, urged that these studies (including a functional observational battery) be extended, refined, and better integrated with general toxicology studies. Others have emphasized the critical contribution of general pharmacology studies to safety in the conduct of early clinical trials (Igarashi *et al.,* 1995; Jorkasky, 1998; Williams, 1990). The limitations of conducting pharmacological measurements in animals under the restrained conditions of a normal toxicology study were acknowledged. Safety pharmacology studies should align with toxicology studies in terms of choice of species and dosage regimen as appropriate.

The value of supporting mechanistic studies, often *in vitro,* was mentioned, but these will not usually be conducted broadly as part of a preclinical screen; however, certain compound classes may trigger particular screens. The *cardiovascular* group drew attention to the use of blood pressure measurement by telemetry, Purkinkje fiber preparations, electrophysiology of myocardial cells, etc. The *hematology* group mentioned flow cytometry and bone marrow culture. While these and other similar experimental tools have unquestionable value for exploration of the mechanism of toxicity, they have not hitherto found application in the prediction of HT. Companies were asked, when a HT had an animal correlate, if it was derived from a standard or a nonstandard toxicology study; only 13/148 cases were from nonstandard studies.

c. The choice of species might also be the subject of more thoughtful consideration. Often studies are conducted, in the dog and the rat, without an open-minded consideration of whether an alternative species might

be better in terms of pharmacodynamics, physiology, biochemistry, metabolism, etc. Guinea pigs are obviously well-established in testing for *hypersensitivity;* the *neurological* group drew attention to the utility of specially trained primates, though again presumably not as a routine preclinical screen. The *liver* and *gastrointestinal* groups suggested that animal disease models could be put to better use. The human diseases being treated may increase patient susceptibility to a, e.g., through increased gastrointestinal permeability which would not be reflected in animals with normal gastrointestinal reserve.

d. Several groups urged the more imaginative use of biomarkers. These could perhaps reveal, on the one hand, hitherto undetected changes in animals and, on the other, earlier detection of HTs. The biomarkers could be different in animals and in humans. Examples included troponin T and CK-MB in the *cardiovascular* area and $\alpha$2-antitrypsin in the *gastrointestinal* area. Along the same lines, the *hypersensitivity* group mentioned the underestimated value of lymph node assays and that systematic weighing of immune/lymphoid organs in animal studies often provides a first indication of a disturbance to the immune system. The *hematologic* group drew attention to the potential for wider exploitation of newer technique such as flow cytometry and bone marrow culture.

e. The *hypersensitivity* group pointed out that since the present design of animal preclinical toxicology studies has little scope for assessment of immunological endpoints, the poor human–animal correlation for this HT category could only be rectified by the routine addition of tests or models aimed at the detection of systemic or cutaneous hypersensitivity and, where relevant, phototoxicity, in relevant species. This group also acknowledged that whereas phototoxicity testing in guinea pigs is fairly reliable as an indicator of human hazard, other test systems for hypersensitivity are far from satisfactory and there is need for urgent research in this area.

f. One way to diminish dependence on the animal to human extrapolation process is to work directly with human tissue. The *liver* group noted the potential value of human liver slices and other *ex vivo* or *in vitro* preparations to obtain of information on metabolic transformation of test compounds and, conversely, of potential effects of the test compound on the liver.

g. Several groups speculated, without being specific, on the possible future use of molecular biological techniques, such as gene expression profiling proteomics, the use of gene chips, etc. The main value may be in identifying in advance individuals with intrinsic susceptibility to various HTs.

## DISCUSSION

This study did not attempt to assess the predictability of preclinical experimental data to humans. What it evaluated was the concordance between adverse findings in clinical data with data which had been generated in experimental animals (preclinical toxicology).

This HESI Workshop and collaborative project is unique (to our knowledge) in magnitude of the database and scope of project. This is an initial step to develop a quantitative understanding of concordance of animal target organ toxicology and manifest HT associated with pharmaceutical development. The intent of this project at its inception was to relate the value of preclinical testing models and methods to identify important HT, which by definition is "relevant" toxicity. This approach provides useful perspective for the types of HT evaluated. It is recognized to be limited by not being able to fully explore all aspects of "predictivity" of HT (see Fig. 1).

No restrictions were placed on the time period from which qualifying data sets could be submitted. Indeed, the inclusive years of data collection for the full database are unknown. Factors influencing the individual data submissions include the refinement of protocols in recent years, unavailability of certain types of data and endpoints from earlier studies, and the development of GLP protocols which might impact data quality. Other than limiting the availability of certain types of information such as data on metabolites and toxicokinetics, it is unknown whether the unevenness of study designs over time might have other influences on the outcome of the database analysis.

A significant message from this survey is that the two HTs with the poorest correlation with animal studies (liver and hypersensitivity/cutaneous reactions) were also the two HTs that led most often to termination of clinical development. This highlights the need for progress in these two areas. The way forward must surely be to increase investigations of mechanism; each occurrence of any unpredicted HT (not just liver or skin) should be followed by investigation of the mechanism involved which, in turn, should lead to a search for a nonclinical predictive model.

The results of this survey and the workshop breakout group discussions have identified several key findings and have also revealed several areas for additional evaluation to pursue in a future project. The main finding of this study is the true positive concordance rate of over 71% for comparable target organs in animal toxicity studies for identified HTs. In addition the survey supports the utility and relevance of studies of up to 1 month in duration with target organ toxicity alerts seen in over 90% of cases. Prior to ICH recommendations on parity of preclinical versus clinical dosing duration, many companies performed 1-month studies before entry into Phase I.

**TABLE 6**
**Distribution of Therapeutic Classes (%)**

| | This database | World scene | |
|---|---|---|---|
| | | In development Dec. 1997[a] | Marketed 1989–1998[b] |
| Anticancer | 9 | 16 | 8 |
| Anti-infective | 14 | 9 | 16 |
| Cardiovascular | 11 | 15 | 21 |
| Endocrine | 7 | 10 | 8 |
| Gastrointestinal | 6 | 3 | 3 |
| Immunological | 1 | 7[c] | 13[c] |
| Neurological | 21 | 24 | 16 |
| Respiratory | 9 | 6 | 4 |
| Other | 22 | 9 | 11 |

[a] *Source:* Ashton *et al*. (1998) (data from 42 companies).
[b] From CMR (1999).
[c] Includes anti-inflammatory therapies.

Since 39% of the HTs described in this study were first observed in short-term duration Phase I clinical studies, including 28% seen following single-dose administration to man, it may be important for a future survey to determine the exposure ratio (duration and therapeutic index) between animal study findings and HT occurrence. This may help to explain why compounds were progressed in the clinic despite preclinical evidence of potential toxicity. Additionally, in a future prospective survey it would be useful to identify the duration of animal studies required to identify all HTs of a specific type and especially the time to onset of those HTs observed in Phase II and III clinical trials. Of course there remains important value to the conduct of longer duration preclinical studies including changes in NOELs over time, progression of target organ toxicities with chronic administration, and evaluation of processes leading to carcinogenesis. These aspects were not covered in the current survey.

A remaining need, and shortcoming of this survey—as pointed out in the Introduction—is that the design did not include "false positive" and "true negative" outcomes to determine the discriminating value of prospective preclinical toxicity biomarker signals to predict HTs. A more complete evaluation of this predictivity aspect will be an important part of a future prospective survey.

A question raised at the Workshop concerned to what extent this database was representative of the range of drug types under development and marketed worldwide. If it were not, caution would be called for in drawing general conclusions about the reliability of animal models. A comparison of the distribution of therapeutic classes in the present database with that of drugs in development and marketed in recent years is shown in Table 6. This shows that relative to drugs in development, the HESI database is somewhat overrepresented by anti-infective agents and underrepresented by anti-cancer agents; in comparison with marketed drugs, there is some overrepresentation of neurologic and respiratory classes and some underrepresentation of cardiovascular and immunological classes. Although unlikely, it cannot be excluded that this may be due, in part, to the varied vintage of the compounds submitted. None of the discrepancies would be sufficient to seriously distort the overall conclusions of this survey.

The Workshop concluded that the project had been of real value in bringing together scientists from the pharmaceutical industry, academia, and government regulatory agencies to discuss the strengths and weaknesses of the current nonclinical toxicology strategies. It was agreed that it was desirable that the project should proceed to a second phase, essentially to broaden and extend the present database to ideally add measures of true negative and false positive rates for preclinical toxicology. Lessons learned in this first phase, discussed above in response to Q.1, would be applied to make the next data collection more informative and open to additional questioning. Additionally, in a future exercise one would be assured that the data submitted were representative of current pharmaceutical research activity. Other companies wishing to participate are invited to contact Karluss Thomas, one of the authors of this paper.

## ACKNOWLEDGMENTS

We recognize the contributions of the cochairs of the six breakout groups to the HESI Workshop in providing organization and leadership at the workshop sessions. They are Dr. Louis Cantilena, Armed Forces Institute of Pathology; Dr. Gabre-Mariam Mesfin, Pharmacia & Upjohn; Dr. Ralph Heywood, Independent Consultant; Dr. Fred Radzialowski, Monsanto/Searle Laboratories; Dr. Anthony Dayan, University of London; Dr. Nasir Khan, Monsanto/Searle Laboratories; Dr. Neal Shear, Sunnybrook Health Science Center; Dr. Joseph Vos, National Institute of Public Health and the Environment (The Netherlands); Dr. Francois Ballet, Rhone Poulenc Rorer; Dr. Paul Watkins, University of Michigan Medical School; Dr. Marc Bonnefoi, Rhone-Poulenc Rorer; and Dr. Pamela Cyrus, Bayer Corporation. Summary reports of each breakout session topic are available upon request from Karluss Thomas (HESI). Acknowledgments are also extended to the management of all sponsor companies for providing input into the database used in this survey. Participating companies included Abbott Laboratories, Bayer Corporation, Boehringer-Ingelheim Pharmaceuticals Inc., Eli Lilly & Company, Janssen Research Foundation, Monsanto/Searle Laboratories, Novartis Pharmaceutical Corporation, Pfizer Pharmaceuticals, Pharmacia & Upjohn Inc., Rhone-Poulenc Rorer Inc., Sanofi Pharmaceuticals Inc., and Zeneca Pharmaceuticals. Additionally, acknowledgments also are extended to all HESI Workshop participants for providing input and critique of the database and the survey instrument, to improve the planned Phase 2 survey.

## REFERENCES

Ashton, G. A., Lobo, L. I., Griffiths, S. A., and McAuslane, N. (1998). Activities of the international pharmaceutical industry in 1997: Pharmaceutical investment and output. *Centre Med. Res. CMR* 98–104R.

Brimblecombe, R. (1990). The importance of retrospective comparisons. In *CMR Workshop—Animal Toxicity Studies: Their Rele-*

*vance for Man* (C. E. Lumley and S. W. Walker, Eds.), pp. 15–19. Quay, Lancaster, UK.

Calabrese, E. J. (1984). Suitability of animal models for predictive toxicology: Theoretical and practical considerations. *Drug Metab. Rev.* **15,** 505–523.

Calabrese, E. J. (1987). *Principles of Animal Extrapolation*. Wiley, New York.

CMR (1999). *CMR News,* Vol. 17. Centre for Medicines Research International, Epsom, Surrey, UK.

Fletcher, A. P. (1987). Drug safety tests and subsequent clinical experience. *J. R. Soc. Med.* **71,** 693–696.

Garratini, S. (1985). Toxic effects of chemicals: difficulties in extrapolating data from animals to man. *Annu. Rev. Toxicol. Pharmacol.* **16,** 1–29.

Heywoood, R. (1990). Clinical toxicity—Could it have been predicted? Post-marketing experience. In *Animal Toxicity Studies: Their Relevance for Man* (C. E. Lumley and S. W. Walker, Eds.), pp. 57–67. Quay, Lancaster, UK.

Igarashi, T. (1994). The duration of toxicity studies required to support repeated dosing in clinical investigation—A toxicologists opinion. In *CMR Workshop: The Timing of Toxicological Studies to Support Clinical Trials* (C. Parkinson, N. McAuslane, C. Lumley, and S. R. Walker, Eds.), pp. 67–74. Kluwer, Boston/UK.

Igarashi,T., Nakane, S., and Kitagawa, T. (1995). Predictability of clinical adverse reactions of drugs by general pharmacology studies. *J. Toxicol. Sci.* **20,** 77–92.

Jorkasky, D. K. (1998). What does the clinician want to know from the toxicologist? *Toxicol. Lett.* **102–103,** 539–543.

JPMA (1994). Seiyakukyo data No. 65. Translation of a report from the Japanese Pharmaceutical Manufacturers Association, provided by Dr. T. Igarashi.

Lawrence, D. R., McLean, A. E. M., and Weatheral, M. (1984). *Safety Testing of New Drugs*. Academic Press, New York.

Litchfield, J. T. (1962). Evaluation of the safety of new drugs by means of tests in animals. *Clin. Pharmacol. Ther.* **3,** 665–672.

Lumley, C. (1990). Clinical toxicity: Could it have been prevented? Premarketing experience. In *CMR Workshop—Animal Toxicity Studies: Their Relevance for Man* (C. E. Lumley, and S. W. Walker, Eds.), pp. 49–56. Quay, Lancaster, UK.

Lumley. C. E., and Walker, S. R. (Eds.) (1990). *CMR Workshop—Animal Toxicity Studies: Their Relevance for Man*. Quay, Lancaster, UK.

McLean, A. (1990). In *CMR Workshop—Animal Toxicity Studies: Their Relevance for Man* (C. E. Lumley and S. W. Walker, Eds.), pp. 79–82. Quay, Lancaster, UK.

National Technical Information Service (1999). *COSTART: Coding symbols for Thesaurus of Adverse Reaction Items,* Version 5. National Technical Information Service, Springfield, VA.

Olson, H., Betton, G., Stritar, J., and Robinson, D. (1998). The predictivity of the toxicity of pharmaceuticals in humans from animal data—An interim assessment. *Toxicol. Lett.* **102–103,** 535–538.

Oser. B. L. (1981). The rat as a model for human toxicological evaluation. *J. Toxicol. Environ. Health* **8,** 521–642.

Owens, A. H. (1962). Predicting anticancer drug effects in man from laboratory animal studies. *J. Chron. Dis.* **15,** 223–228.

Parkinson, C., McAuslane, N., Lumley, C., and Walker, S. R. (Eds.) (1994). *CMR Workshop: The Timing of Toxicological Studies to Support Clinical Trials*. Kluwer, Boston/UK.

Prentis, R. A., *et al.* (1988). Pharmaceutical innovation by the seven UK-owned pharmaceutical companies. *Br. J. Clin. Pharmacol.* **25,** 387–396.

Rozencweig, M., *et al.* (1981). Animal toxicology for early clinical trials with anticancer agents. *Cancer Clin. Trials* **4,** 21–28.

Schein, *et al.* (1970). The evaluation of anticancer drugs in dogs and monkeys for the prediction of qualitative toxicities in man. *Clin. Pharmacol. Ther.* **11,** 3–40.

Spriet-Pourra, C., and Auriche, M. (Eds.) (1994). *SCRIP Reports*. PJB, New York.

Williams, P. D. (1990). The role of pharmacological profiling in safety assessment. *Regul. Toxicol. Pharmacol.* **12,** 238–252.

Zbinden, G. (1991). Predictive value of animal studies in toxicology. *Regul. Tox. Pharm.* **14,** 167–177.

Zbinden, G. (1993). The concept of multi-species testing in industrial toxicology. *Regul. Toxicol. Pharmacol.* **17,** 85–94.